113191

NR-238-001
Contract Nonr-609(02)

Edwards Street Laboratory
Yale University
New Haven, Connecticut

ESL Technical Report No. 21
(ESL:590:Ser 16)
31 December 1953

The Effect of Errors of Curvilinear Coordinates on the Error of Position of a Point

Donald Foster

Copy 7

NR-238-001
Contract Nonr-609(02)

Edwards Street Laboratory
Yale University
New Haven, Connecticut

ESL Technical Report No. 21
(ESL:590:Ser 16)
31 December 1953

The Effect of Errors of Curvilinear Coordinates on the Error of Position of a Point

Donald Foster

Abstract

The position of a point on a surface, as determined by measurement, is subject to errors of observation. The direction and distance of the measured point from the true point depends not only on the errors of the measured coordinates but also on the properties of the two families of lines of which the measured coordinates are the independent parameters. Practical considerations generally preclude the use of measured coordinates for which the error of position depends only on the errors of observation, or for which the effect of the coordinate system is easily understood.

When the distribution functions of the errors are known, it is possible to construct around any point representing the true point on the surface a closed curve along which the probability of occurrence of a corresponding measured point is uniform, and inside which there is a specified probability that a measured point will fall. Accordingly, the same closed curve may be drawn around a measured point to represent the size and shape of the smallest region inside which, with the specified probability, the (unknown) true point will lie.

Details are worked out for two kinds of error functions -- the normal law, and errors of limited size. For the former,

the paper shows how to calculate, for each position of a measured point on the surface, the boundary inside which there is a specified probability that the true point will lie. In the latter case, equations are given which define the smallest region inside which the true point will certainly be found.

While the analysis has been restricted to the plane, the ideas are applicable to other surfaces and to three dimensions.

Introduction

The coordinate system employed to find the position of a point by physical measurement is usually determined by necessity, convenience, or accuracy in making the measurement. The choice must often be mathematically inconvenient. In two dimensions each measured coordinate is a parameter of one of two families of loci. A pair of coordinates determine a particular locus from each of the two different families; and these two curves intersect to determine a point.

Consider the two coordinate curves which pass through the actual point and another pair of curves whose parameters are erroneous measured coordinates. These four curves form what may be called a curvilinear parallelogram, with the point at one vertex and the erroneous point at the opposite vertex. The distance between these two points depends on the sides of the parallelogram and the angles at its vertices. These elements in turn depend on the size and relative signs of the errors, the nature of the coordinate system, and the values of the coordinates at the point. The relations between an error of position and the errors of observation from which it arises are much more complicated in general than they are in the cartesian system or even in orthogonal curvilinear systems. Since it is often impossible to avoid using complicated coordinate systems, it is necessary to understand these matters if one is to make a choice of measuring tools and their

locations which will satisfy the requirements.

Probability considerations

Let U and V be the curvilinear coordinates of a point in the X,Y plane; and let U + u, V + v be erroneous measured values of the same pair of coordinates. Assume that U and V have the same physical dimensions. Let their errors, u and v, be of normal distribution and of known precision. The probability that an error in U lies between u and u + du while the corresponding error in V lies between v and v + dv is then the function

$$\frac{kk'}{\pi} e^{-(k^2u^2 + k'^2v^2)} du\,dv \qquad (1)$$

where the k's are experimentally determined moduli of precision. The set of all points for which this probability is uniform are subject to the condition

$$k^2u^2 + k'^2v^2 = h \qquad (2)$$

The parameter, h, may be determined by substitution in (2) of any particular pair of values of u and v. The corresponding value of the probability, (1), is

$$\frac{kk'}{\pi} e^{-h} du\,dv \qquad (3)$$

and this value is uniform over the band between the two ellipses whose parameters are h and h + dh. Since the area of the ellipse, (2), is

$$A = \pi \frac{\sqrt{h}}{k} \frac{\sqrt{h}}{k'}$$

the area of the region between the ellipses is

$$dA = \frac{\pi}{kk'} dh$$

Changing the element of area in (3) from dudv to dh, we find the probability of a point being inside any particular ellipse is

$$P = \int_0^h e^{-h} dh = 1 - e^{-h} \tag{4}$$

If k = k', the ellipses, (2), become circles

$$u^2 + v^2 = r^2 \; ; \; \left(r^2 \equiv \frac{h}{k^2} \right) \tag{5}$$

for which

$$P = 1 - e^{-k^2 r^2} \tag{6}$$

Imagine that we have made a very large number, N, of pairs of measurements (U + u , V + v) of the coordinates (U,V) of a single point. Suppose all these measured points to be plotted in the U,V plane and also in the X,Y plane. In terms of the N plotted points in the U,V plane, the above equations have the following interpretation. The density of points per

unit area is

$$N\frac{kk'}{\pi}\exp(-k^2u^2-k'^2v^2)$$

Over the narrow band included between two ellipses of the family,

$$k^2u^2+k'^2v^2=h$$

whose parameters are h and h + dh, the density of points is uniform and the number of points included in the band is $Ne^{-h}dh$. Suppose we divide the space around (U,V) into a number of such adjacent elliptical bands by constructing several concentric ellipses of the family (2). The probability that a single measured point will fall outside one of these ellipses is a simple function of the parameter h, as we have seen in (4). The picture drawn here in the U,V plane is exactly the same as we should have in the X,Y plane, had we been measuring X and Y directly instead of indirectly through U and V which are functions of X and Y. The question of practical interest is: what does this plot in the U,V plane look like when it is mapped into the X,Y plane? Under any point transformation X = X(U,V), Y = Y(U,V), we should expect the ellipses (or circles, when k = k') to map into some kind of closed curves with the images of the same points as before lying between them. In other words, after the mapping, the probability of a single measured point falling outside a given closed curve in the X,Y plane is the same as that of its falling outside the image of this curve in the U,V

plane. It will be shown that, subject to very mild restrictions, small ellipses (or circles) in one plane always map into ellipses in the other plane. The size, orientation, and eccentricity of the mapped ellipse generally vary from point to point in the plane. We now proceed to examine the nature of these transformations.

Equations for the mapping of small regions

Let the measured coordinates (U,V) be related to the cartesian (X,Y) through

$$U = U(X,Y) \qquad V = V(X,Y) \tag{7}$$

or the inverse

$$X = X(U,V) \qquad Y = Y(U,V) \tag{8}$$

in which the functions are all supposed to be real. Either (7) or (8) provide all necessary information for mapping one plane into the other; but since we are interested only in mapping separate small regions we may take advantage of the fact that the transformation for small regions is generally linear. In place of (7) or (8), which may be of any form, we may substitute linear transformations, the elements of whose matrices are the partial derivatives of (7) or (8) evaluated at some fixed point in the small region being mapped.

If X and Y may be expanded in Taylor series in the vicinity of a point (U_o, V_o), we have

$$x = \frac{\partial X}{\partial U}u + \frac{\partial X}{\partial V}v + \frac{1}{2}\left(\frac{\partial^2 X}{\partial U^2}u^2 + 2\frac{\partial^2 X}{\partial U \partial V}uv + \frac{\partial^2 X}{\partial V^2}v^2\right) + \cdots$$

$$y = \frac{\partial Y}{\partial U}u + \frac{\partial Y}{\partial V}v + \frac{1}{2}\left(\frac{\partial^2 Y}{\partial U^2}u^2 + 2\frac{\partial^2 Y}{\partial U \partial V}uv + \frac{\partial^2 Y}{\partial V^2}v^2\right) + \cdots \tag{9}$$

in which $x = X - X_o$, $y = Y - Y_o$, $u = U - U_o$, $v = V - V_o$ are finite increments which we may identify with the errors under discussion. In (9), the partial derivatives are to be evaluated at (U_o, V_o). With proper restrictions as to continuity and size of the region we may therefore employ the equations

$$\begin{aligned} x &= \frac{\partial X}{\partial U}u + \frac{\partial X}{\partial V}v \\ y &= \frac{\partial Y}{\partial U}u + \frac{\partial Y}{\partial V}v \end{aligned} \tag{10}$$

and

$$\begin{aligned} u &= \frac{\partial U}{\partial X}x + \frac{\partial U}{\partial Y}y \\ v &= \frac{\partial V}{\partial X}x + \frac{\partial V}{\partial Y}y \end{aligned} \tag{11}$$

in which the coefficients $\frac{\partial X}{\partial U}$, etc. are constants.

In order to shorten the notation and emphasize the linear character of the approximations, we re-write (10) and (11) as

$$\begin{aligned} x &= a'_{11}u + a'_{12}v \\ y &= a'_{21}u + a'_{22}v \end{aligned} \tag{12}$$

and

$$u = a_{11} x + a_{12} y$$
$$v = a_{21} x + a_{22} y \qquad (13)$$

in which all the a_{ij} are real.

While we may use these approximations freely, because they are generally accurate, it is true that one could encounter circumstances in which the parameter h is so large and the successive partial derivatives of such values at particular points of the field that (12) and (13) would be inaccurate. However, the main purport of what we have to say about the probabilities associated with the images in the X,Y plane of the ellipses (2) in the U,V planes is unaffected by whether we can employ (12) and (13) or must use (7) and (8).

The important fact is that, given any probability functions, we can calculate the contours of constant probability in the U,V plane and then draw the corresponding contours of the same probabilities in the X,Y plane. For monotonic decreasing probability functions, the latter show the size and shape of the smallest regions inside which there is a specified probability that a measured point will occur. It is evident that this can always be done, regardless of the shape of the monotonic error functions, or what special

devices are employed to effect the transformation. Thus the basic ideas of this discussion are independent of any particular assumptions such as (1) or the use of linear transformations.

Elements of the ellipses in the X,Y plane

The ellipse (2), in the U,V plane, becomes in the X,Y plane

$$1 = \frac{1}{h}(k^2u^2 + k'^2v^2) = \frac{1}{h}\Big[(k^2a_{11}^2 + k'^2a_{21}^2)x^2 + 2(k^2a_{11}a_{12} + k'^2a_{21}a_{22})xy + (k^2a_{12}^2 + k'^2a_{22}^2)y^2\Big] \tag{15}$$

The discriminant of the quadratic form on the right is

$$D \equiv -4k^2k'^2\Delta^2 < 0\ ; \quad \left(\Delta \equiv \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}\right) \tag{16}$$

which shows that an ellipse transforms into an ellipse. Obviously this is true regardless of the direction of transformation.

Let us write (15) in the more compact form

$$\frac{1}{h}(Ex^2 + 2Fxy + Gy^2) = 1 \quad * \tag{15'}$$

---

* Note that E, F, and G as here defined are not identical with the familiar E, F and G of differential geometry except when k = k' = 1.

and refer it to new coordinates x',y' so that it takes the form

$$\frac{1}{k}\left(\frac{x'^2}{a^2}+\frac{y'^2}{b^2}\right)=1 \tag{17}$$

This we can do by means of a rotation about the point X_o, Y_o -

$$\begin{aligned} x &= x'\cos\vartheta - y'\sin\vartheta \\ y &= x'\sin\vartheta + y'\cos\vartheta \end{aligned} \tag{18}$$

such that

$$\tan 2\vartheta = \frac{2\mathcal{F}}{\mathcal{E}-\mathcal{G}}; \quad (0 \leq 2\vartheta \leq \pi) \tag{19}$$

As E and G are positive, we find in the relations

$$\cos 2\vartheta = \frac{\mathcal{E}-\mathcal{G}}{\pm\mathcal{R}} \qquad \sin 2\vartheta = \frac{2\mathcal{F}}{\pm\mathcal{R}}$$

$$\cos\vartheta = \left\{\frac{\mathcal{R}\pm(\mathcal{E}-\mathcal{G})}{2\mathcal{R}}\right\}^{1/2} \qquad \sin\vartheta = \left\{\frac{\mathcal{R}\mp(\mathcal{E}-\mathcal{G})}{2\mathcal{R}}\right\}^{1/2} \tag{20}$$

where $\mathcal{R} \equiv +\sqrt{(\mathcal{E}-\mathcal{G})^2+4\mathcal{F}^2} = +\sqrt{(\mathcal{E}+\mathcal{G})^2+D}$

$$= +\sqrt{(\mathcal{E}+\mathcal{G})^2 - 4k^2k'^2\Delta^2}$$

that the upper sign is to be taken when F $>$ 0 and the lower sign when F $<$ 0. The numbers a and b in (17) are found to be given by

$$a^2 = \frac{2}{\mathcal{E}+\mathcal{G}\pm\mathcal{K}} \quad ; \quad b^2 = \frac{2}{\mathcal{E}+\mathcal{G}\mp\mathcal{K}} \tag{21}$$

where the choice of signs depends on the sign of F in the same way as before.

We note that

$$R = 0$$

implies that

$$E = G \quad ; \quad F = 0$$

These are the conditions for the ellipse to be a circle. The eccentricity of the ellipse is

$$e = \left(\frac{2\mathcal{R}}{\mathcal{E}+\mathcal{G}+\mathcal{R}}\right)^{1/2} \tag{22}$$

If the conditions $\mathcal{E} \neq \mathcal{G}$; $\mathcal{F} = 0$

are satisifed the locus is an ellipse whose axes are parallel to the axes of X and Y and whose eccentricity is

$$e = \left(1 - \frac{\mathcal{E}}{\mathcal{G}}\right)^{1/2}, \ (\mathcal{E} < \mathcal{G}) \quad ; \quad e = \left(1 - \frac{\mathcal{G}}{\mathcal{E}}\right)^{1/2}, \ (\mathcal{E} > \mathcal{G})$$

The major and minor semi-axes in the general case are

$$\left(\frac{2h}{\mathcal{E}+\mathcal{G}-\mathcal{R}}\right)^{1/2} \quad \text{and} \quad \left(\frac{2h}{\mathcal{E}+\mathcal{G}+\mathcal{K}}\right)^{1/2} \tag{23}$$

and when F = 0, they are

$$\left(\frac{h}{\mathcal{G}}\right)^{1/2} \text{and} \left(\frac{h}{\mathcal{E}}\right)^{1/2}, \ (\mathcal{E} > \mathcal{G})$$

or

$$\left(\frac{h}{\mathcal{E}}\right)^{1/2} \text{ and } \left(\frac{h}{\mathcal{G}}\right)^{1/2}, \quad (\mathcal{E} < \mathcal{G})$$

The chief results of this section are the expressions for major and minor semi-axes, (23), and the direction of the x' axis, (20). The rule of signs implies that when $F > 0$ the major axis lies in the second quadrant, and when $F < 0$ it is in the first quadrant.

Bounded errors

Instead of following the normal distribution, it may be found that errors of absolute value greater than a certain small number do not occur. For example, a set of measurements of an angle by an observer using a certain sextant may show that the chance of an error greater than two minutes is negligible. If this is true, we may employ as the smallest contour of unit probability in the U,V plane the rectangle formed by the four lines

$$\begin{aligned} u &= u_0 \qquad & u &= -u_0 \\ v &= v_0 \qquad & v &= -v_0 \end{aligned} \tag{24}$$

which maps into the oblique parallelogram bounded by the lines

$$\left.\begin{aligned} u &= a_{11}x + a_{12}y \\ v &= a_{21}x + a_{22}y \end{aligned}\right\}, \quad (u = \pm u_0, v = \pm v_0) \tag{25}$$

The coordinates (x,y) of the four vertices of the parallelogram, referred to (X_0, Y_0) as origin, are given by the inverse transformation, (12),

$$\left.\begin{aligned} x &= a'_{11}u + a'_{12}v \\ y &= a'_{21}u + a'_{22}v \end{aligned}\right\}, \quad (u = \pm u_0,\ v = \pm v_0) \tag{26}$$

and the squares of the semi-diagonals are given by the two values assumed by

$$\begin{aligned} x^2+y^2 &= (a'^2_{11} + a'^2_{21})u^2 + 2(a'_{11}a'_{12} + a'_{21}a'_{22})uv + (a'^2_{12} + a'^2_{22})v^2 \\ &= \mathcal{E}'u^2 + 2\mathcal{F}'uv + \mathcal{G}'v^2, \ \left(\begin{aligned} &\text{where } uv = u_0v_0 > 0 \\ &\text{and } uv = u_0v_0 < 0 \end{aligned}\right) \end{aligned} \tag{27}$$*

With proper attention to sign, the angles at the vertices of the parallelogram are given by

$$\omega = \cos^{-1}\frac{\mathcal{F}'}{\sqrt{\mathcal{E}'\mathcal{G}'}}$$

and the sides are $u\sqrt{\mathcal{E}'}$ and $v\sqrt{\mathcal{G}'}$.

The slopes of the diagonals are given by

$$\left.\frac{y}{x}\right]_{\substack{u=u_0\\v=v_0}} = \left.\frac{y}{x}\right]_{\substack{u=-u_0\\v=-v_0}} = \frac{a'_{21}u_0 + a'_{22}v_0}{a'_{11}u_0 + a'_{12}v_0} \tag{28}$$

---

* In this case E', F', and G' are identical with the coefficients of the first fundamental quadratic form of differential geometry, $dS^2 = dX^2 + dY^2 = E'dU^2 + 2F'dUdV + G'dV^2$.

and

$$\left.\frac{y}{x}\right]_{\substack{u=u_o\\ v=-v_o}} = \left.\frac{y}{x}\right]_{\substack{u=-u_o\\ v=v_o}} = \frac{a_{21}'u_o - a_{22}'v_o}{a_{11}'u_o - a_{12}'v_o}$$

The major diagonal is that for which the product of u_o and v_o has the same sign as F'. Along the locus of F' = 0 the parallelograms become rectangles.

If $u_o = v_o$, the lengths squared of the semi-diagonals become

$$x^2+y^2 = (\mathcal{E}' \pm 2\mathcal{F}' + \mathcal{G}')\,u_o^2$$

and their slopes are

$$\frac{a_{21}'+a_{22}'}{a_{11}'+a_{12}'} \; ; \; \frac{a_{21}'-a_{22}'}{a_{11}'-a_{12}'}$$

The assumption of square form for the contour P = 1, in the U,V plane is not inconsistent with $u_o \neq v_o$ because we may take the side of the square equal to the longer side of the rectangle. All the points then lie inside the square. The only loss is that the square is not the smallest contour for which P = 1.

<u>Graphical representation of the information</u>

The assumption of bounded errors, while crude, may sometimes be near enough to the truth. When it can be used together with the condition $u_o = v_o$ we may describe the manner in which the transformation propagates errors by the

normalized major semi-diagonal

$$\left(\frac{x^2+y^2}{u_0^2}\right)^{1/2} = \left(\mathcal{E}' + |2\mathcal{F}'| + \mathcal{G}'\right)^{1/2}$$

which is the quotient of the longer diagonal of the parallelogram divided by the side of the square. If the contours of the surface

$$f(X,Y) = \mathcal{E}' + |2\mathcal{F}'| + \mathcal{G}'$$

are projected on the X,Y plane they may be regarded as loci of points for which the circle circumscribed around the parallelogram is of constant diameter. In other words they are loci of constant maximum possible error. The disadvantage of this simple representation is that a portion of the information has been thrown away. We have discarded the information about the two directions in which, for $u_o = v_o$, the largest measured errors may yield their largest and smallest displacements in the X,Y plane.

A more complete way of presenting the facts would be to plot at intervals along each contour of the normalized major semi-diagonal, $\left(\mathcal{E}' + |2\mathcal{F}'| + \mathcal{G}'\right)^{1/2}$, a vector showing the direction and magnitude of the normalized minor semi-diagonal and a vector showing the direction of the major semi-diagonal.

In the case of the normal law of errors the elements of the ellipses for, say $P = p_1$, can be represented by plotting small L shaped pairs of vectors indicating by their length and directions the major and minor axes of the ellipses as

they vary over the X,Y plane.

Measured coordinates not of the same physical dimensions

So far in the discussion we have supposed that U and V are of the same kind. If they are not of the same kind as, for example, in ordinary polar coordinates, there are no difficulties in applying the theory for normal error distributions. However, since the moduli of precision are then of different dimensions, it is futile to discuss a particular case in which k = k'.

In the case of bounded errors the preceding treatment is still valid, the only restriction being that to compare the magnitudes of u_o and v_o is without meaning.

Example

For the sake of brevity we illustrate the theory with an example of such simplicity that all the conclusions are self-evident. Consider the case of ordinary polar coordinates. Having no further need to represent the differential of an error, and not wishing to use unfamiliar symbols for the polar coordinates, we write the transformation as

$$r = (x^2 + y^2)^{1/2}$$

$$\vartheta = \tan^{-1}\frac{y}{x}$$

and the relations between the errors, written as differentials, as

$$d r = \cos\vartheta \cdot dx + \sin\vartheta \cdot dy$$

$$d\vartheta = -\frac{\sin\vartheta}{r} \cdot dx + \frac{\cos\vartheta}{r} \cdot dy$$

The equation of a normal probability contour is

$$\begin{aligned} h &= k^2 dr^2 + k'^2 d\vartheta^2 \\ &= \left(k^2\cos^2\vartheta + \frac{k'^2}{r^2}\sin^2\vartheta\right) dx^2 + 2\cdot\tfrac{1}{2}\sin 2\vartheta \left(k^2 - \frac{k'^2}{r^2}\right) dx\,dy \\ &\quad + \left(k^2\sin^2\vartheta + \frac{k'^2}{r^2}\cos^2\vartheta\right) dy^2 \end{aligned}$$

Hence

$$\mathcal{E} + \mathcal{G} = k^2 + \frac{k'^2}{r^2}\ ; \quad \mathcal{R} = k^2 - \frac{k'^2}{r^2}$$

The semi-axes of the ellipse are therefore

$$\frac{\sqrt{h}}{k}\ , \quad \frac{r\sqrt{h}}{k'}$$

Note that the former is independent of position; and the latter is proportional to r and independent of ϑ.

The orientation of the ellipse in the x,y plane is given by

$$\tan 2\phi = \frac{2\mathcal{F}}{\mathcal{E} - \mathcal{G}} = \frac{\sin 2\vartheta \cdot \left(k^2 - \frac{k'^2}{r^2}\right)}{\left(k^2 - \frac{k'^2}{r^2}\right)\left(\cos^2\vartheta - \sin^2\vartheta\right)} = \tan 2\vartheta$$

In other words, one or the other of the axes of the ellipse is always parallel to the radius The ellipses become circles

for any point on the circle

$$r^2 = \left(\frac{k'}{k}\right)^2$$

Examining the sign of F, we find that inside this circle F is negative in the first and third quadrants and positive in the second and fourth. Outside the circle, F is positive in the first and third, and negative in the second and fourth quadrants.

We conclude that the length of the semi-axis which is in the radial direction is uniform and equal to $\sqrt{k}/k$.

The other semi-axis is equal to $r\sqrt{k}/k'$.

When r = k'/k the ellipses are circles.

We next replace the normal law of errors with the assumption of bounded errors and calculate the lengths of the diagonals of the parallelograms in the X,Y plane and their inclinations. The lengths of the semi-axes are given by

$$ds^2 = dx^2 + dy^2 = (\cos\vartheta \cdot dr_0 - r\sin\vartheta \cdot d\vartheta_0)^2 + (\sin\vartheta \cdot dr_0 + r\cos\vartheta \cdot d\vartheta_0)^2$$
$$= dr_0^2 + r^2 d\vartheta_0^2$$

Since there is no term in $dr d\vartheta$, the parallelograms are everywhere rectangles. The slopes of their diagonals are

$$m = \frac{\sin\vartheta \cdot dr_0 \pm r\cos\vartheta \cdot d\vartheta_0}{\cos\vartheta \cdot dr_0 \mp r\sin\vartheta \cdot d\vartheta_0}$$

$$= \tan\left\{\vartheta \pm \tan^{-1}\left(\frac{r\, d\vartheta_0}{dr_0}\right)\right\}$$

These results, in such an extremely simple example, are obvious.

Donald Foster

Donald Foster